Panel Final Comments – Project Report – *Development of a TMDL for the Wanaque Reservoir and Cumulative WLAs/LAs for the Passaic River Watershed* by Najarian Associates, June 2005.

November 29, 2005

At the October 7, 2005 TMDL Advisory Panel Meeting, NJDEP formally requested that the Panel review and comment on Najarian Associates' Final Report entitled "Development of a TMDL for the Wanaque Reservoir and Cumulative WLAs/LAs for the Passaic River Watershed." The Panel has formally commented on earlier drafts of this report.

General Comments

These model results will be useful to the NJDEP in identifying shortcomings in current data and process understanding. In addition, they indicate that a "not to exceed at any time" total phosphorus standard may not be appropriate for this system and that a monthly average or seasonally averaged standard may be more appropriate.

With regard to the steam simulations, one major issue that the Panel has always viewed as a deficiency of this work and has presented earlier comment documents to NJDEP is that the use of a conservative mass-balance approach is not suitable for modeling instream concentrations or daily loadings in the Passaic River watershed system (see the April 14, 2005 comments). Based upon N.J.A.C. 7:15- 7.3 and 7.4, New Jersey does not allow a simple mass balance approach to be used for simulating complex systems (see the April 14, 2005 comments). Additionally, it is not scientifically sound to ignore the instream processes of the system by using the mass-balance approach. Furthermore, monthly discharge self-monitoring data were used to develop a time series of monthly average effluent flows and effluent concentrations that were applied to the model. This approach ignores the daily variability in effluent flows that is critically important in simulating a point source dominated system and equally important when comparing the model predictions to a "not to exceed at any time" total phosphorus standard.

With regard to the reservoir simulations, the reservoir model has difficulties predicting daily total phosphorus fluctuations, but does a reasonable job predicting dissolved oxygen and chlorophyll-a.

The Panel previously recommended a nine-year simulation which would omit the 2002 water year. Year 2002 was considered by Najarian to be the most critical year during the ten-year model simulation period due to prevailing drought conditions and the large-scale diversions needed to restore reservoir capacity. The Panel believes that the 2002 water year may be an anomaly and should be excluded. The recommended nine-year simulation was not conducted as part of this final report.

General Recommendations

The Panel recommends the following:

- Provide a "List of Abbreviations" and a "Glossary."
- Provide more discussion, data, and references with regard to problems (i.e., algal blooms) associated with high phosphorus in the Wanaque Reservoir (e.g., level and frequency).
- Although units may be presented and/or provided in the text, units should be provided in all tables and figures where necessary.
- Clarify the use of lbs/year versus lbs/day for the TMDL (Total Maximum ***Daily*** Load).

River Simulation

Najarian Associates are to be commended for their fine job in presenting the requested simulated versus observed total phosphorus results from their modeling study. They have presented these results as scatter (x/y) plots, and the resultant least squares regression/correlation line shows how well the simulated results match up with the actual data (Figures 3.18 through 3.25). The resultant regression coefficient, r^2, allows for an evaluation of the precision of the model.

It is important to comment on the model's ability to predict the observed data. From Figures 3.2a and 3.2b, it appears the model is doing a fair job at predicting observed concentrations within a certain range. For example, Figure 3.2a shows that the model predicts concentrations from 0.1 to 1.6 mg/l fairly accurately, but the model also predicts peaks as high as 2.5 mg/l without data to confirm these peaks. In examining the correlation analysis associated with this section of stream (Figure 3.18), although the coefficients of determination (r^2) appear to be good, the correlation analysis does not go through the origin of the graph, thereby providing a y-intercept value of 0.05 mg/l. This can be interpreted as the model will tend to over-predict total phosphorus concentrations. In fact, by constructing a 1:1 correspondence line through the origin of this graph, one can easily see that the model tends to over-predict because almost all the data are above this line. A 1:1 correspondence line through the origin represents an exact match between observed and predicted solutions.

For the Rockaway River at Pine Brook, the model tends to over-predict at the low total phosphorus concentrations and under-predict at the higher concentrations. Once again, a 1:1 correspondence line through the origin of the graph clearly illustrates this comment. Also, the correlation analysis does not pass through the origin, providing a y-intercept of 0.097 mg/l. Therefore, if the observed total phosphorus concentration was 0 mg/l, the predicted would be 0.097 mg/l. Since the surface water quality criteria for phosphorus are 0.1 mg/l and 0.05 mg/l for streams and lakes, respectively, the model's inability to accurately predict the lower instream phosphorus concentration could result in inappropriate effluent limitations for the point source discharges.

For the Whippany River at Pine Brook, the model is under-predicting the peaks and over-predicting the troughs. From the correlation analysis, as the observed concentrations approach 0 mg/l, the model predictions approach 0.129 mg/l. The low slope (i.e., 0.57) indicates that the model is under-responding to changes in loading or internal transformations of total phosphorus. The correlation has an intercept greater than 0.1, which indicates excess total phosphorus in the model for this location.

For the Passaic River at Two Bridges, the model under-predicts several of the peaks and over-predicts several of the minimum total phosphorus concentrations. From the correlation analysis, as the observed concentrations approach 0 mg/l, the model predictions approach 0.089 mg/l.

For the Ramapo River at Mahwah and Pompton Lakes, the observed instream phosphorus concentrations are relatively low compared to other reaches in the watershed. The model tends to under-predict at Mahwah. Since the scatter in the observed data appears to be significant at Pompton Lakes, it is difficult to compare the model predictions to observations. The low slope (i.e., 0.57) indicates that the model is under-responding to changes in loading or internal transformations of total phosphorus. With r^2 values less than 0.25 (i.e., 0.24), the predictive power of the model at this location is doubtful.

For the Pompton River at Two Bridges, there appears to be inconsistencies when comparing Figure 3.8a to Figure 3.24. Based upon Figure 3.8a, the model under-predicts peak total phosphorus observations over 0.8 mg/l, as well as over-predicts many of the minimum values that were observed. Figure 3.24 does not include any observations over 1.0 mg/l. It appears that the peak observations were not included in this correlation analysis. As is the trend with many of the other reaches, the model tends to over-predict the low concentrations while under-predicting the high concentrations. From the correlation analysis, as the observed concentrations approach 0 mg/l, the model predictions approach 0.12 mg/l. The low slope (i.e., 0.42) indicates that the model is under-responding to changes in loading or internal transformations of total phosphorus. With r^2 values less than 0.25 (i.e., 0.22), the predictive power of the model at this location is doubtful. The correlation has an intercept greater than 0.1 which indicates excess total phosphorus in the model for this location.

For the Passaic River at Little Falls, the model appears to give its best predictions, even though there appears to be scatter in the observed data.

The accuracy of the model can also be evaluated by comparing the regressed lines with the 1:1 line of perfect correlation. This gives an estimate of where the model in most accurate. The point where the regressed lines would intersect the 1:1 line of perfect correlation gives an estimate of where the model is most accurate. This can be obtained from the equation:

$$x = a/(1-b)$$

where x is the total phosphorus concentration where the model is most accurate, a is the regressed intercept and b is the regressed slope. It is possible to obtain a negative value (virtual solution), which may mean that the model has a bias that is manifesting itself as a propagating error as concentration increases. The resultant points of greatest accuracy for each of the figures are presented in Table 1.

Table 1

Figure	Total Phosphorus (mg/l)
3.18	-0.709
3.19	0.289
3.20	0.304
3.21	0.468
3.22	0.0114
3.23	0.108
3.24	0.282
3.25	1.536

It is interesting to note that for the most part, the point of greatest accuracy is substantially greater than the target value of 0.05 mg/l for total phosphorus.

An estimate of the model error at the target value of 0.05 mg/l total phosphorus can be made by using the regression line to obtain a predicted value and then subtracting the target value to obtain a residual. This residual can then be divided by the target to obtain a type of standard error for the model at the target value. These are presented in Table 2.

Table 2

Figure	S.E.(%)
3.18	108
3.19	160
3.20	216
3.21	158
3.22	–14.2
3.23	49.8
3.24	180
3.25	69.8

Based upon this analysis, it can be concluded that the model over-predicts the total phosphorus results on a consistent basis. The model is therefore quite conservative and possibly over-conservative.

In conclusion, we cannot say that the model has been calibrated or validated since the model is a mass-balance approach and does not contain coefficients that can be calibrated. In most stream reaches, the model tends to over-predict the minimum observed total phosphorus concentrations, while it tends to under-predict the

maximum observed concentrations. Since this is a point source dominated system, the model's inability to predict the high and low values may be related to the model's use of average monthly flows and effluent concentrations. The diurnal treatment plant flow patterns may be responsible for producing these peaks and troughs in the observed data. Since the model ignores this diurnal fluctuation of treatment plant flows, it is difficult for the model to simulate the observed data. The model would be better used to predict average monthly instream concentrations or loadings instead of daily values. Since the total phosphorus standard is a "not to exceed at any time" standard, it would be very difficult to use this model to predict reductions in wasteload and load allocations for pollutant sources that achieve the total phosphorus standard.

Reservoir Simulation

The calibration and validation of the reservoir model was presented in "Water Quality Assessment of the Upper Passaic River Watershed and the Wanaque Reservoir," dated September 2000 (Najarian, 2000). Although the Panel was not asked to review this document, below are some comments on the information presented in this report on the LA-WATERS model.

Although LA-WATERS is a model that is appropriate for simulating the Wanaque Reservoir, the accuracy of the model predictions are a function of the quality of the input data. The model requires daily inflow loadings from three sources: 1) the Reservoir's tributary watershed, 2) the Pompton Lakes intake, and 3) the Two Bridges Intake. To understand the model's ability to predict ***daily*** in-lake phosphorus concentrations, the quality of these three inputs must be examined. Although monthly phosphorus data were available on the reservoir tributaries (see Table 2.5), these data were not used to generate model inputs. Instead, unit aerial loading (UAL) coefficients were applied to the watershed land use to develop a time series of daily concentrations. Typically, these UAL coefficients have units of pounds per acre per year and are used to develop gross estimates of nonpoint source loads on a watershed basis. To develop a daily simulation, the loads from nonpoint sources must be linked to rainfall events if there is any hope of predicting daily in-lake concentrations. This involves developing an event mean concentration for each storm event and applying this concentration to the stormwater discharge during the rainfall event. It is unclear if this was the procedure used to develop the time series of daily concentrations needed for model input. Additionally, this time series should be compared to the measured data that were collected by North Jersey District Water Supply Commission in the reservoir tributaries. Excluding these available data is not appropriate unless the data are believed to be suspect, in which case an explanation should have been given on why these data are believed to be bad.

The daily inflow loadings from the Pompton Lakes and Two Bridges intakes was developed using the mass-balance model that was presented earlier in the report. As discussed earlier in these comments, this model typically under-predicts the peak concentrations and over-predicts the minimum concentrations for total phosphorus. The use of this model to develop an accurate daily time series of loading inputs for LA-WATERS is questionable.

Figures 4.10 through 4.15 provide a comparison of model predictions to observed data. From Figure 4.10 it is clear that the model does not predict the daily fluctuations in total phosphorus, and no statistical analysis was provided comparing model predictions to observations. These statistical data were not presented in Najarian 2000. From Figure 4.13, it is clear when comparing total phosphorus predictions at Raymond Dam to observations at Raymond Dam that the model over-predicts the observations made in early 2002, a time period which is later deemed the critical period. The peak observation during this time period is approximately 0.11 mg/l, while the model predicts 0.20 mg/l. In this graph it is unclear why other observations are being compared to computed data at Raymond Dam (i.e., Wanaque River below the Raymond Dam and Dam 4). The model predictions at the surface near West Brook (Figure 4.14) tend to over-predict minimum concentrations while under-predicting the peak concentrations. Once again, the predictions made for early 2002 were higher than the observations. From Figure 4.15, the model tends to under-predict the observed total phosphorus concentrations. Although it is difficult to compare how well the model is predicting without a statistical analysis, the ability of the model to predict daily fluctuations in total phosphorus appears to be limited.

Figures 4.18 through 4.20 illustrate the model's ability to predict dissolved oxygen (DO). Although measured data are limited and cannot be used to examine the daily diurnal fluctuation of DO, the model does a fairly good job predicting DO. From Figures 4.21 and 4.22, the model does a fairly good job predicting chlorophyll-a. It tends to under-predict the peaks in the summer of 2000 and 2001 but does a fairly good job in the other years.

In conclusion, even though the model tends to have difficulties in predicting the daily total phosphorus fluctuations at maximum and minimum concentrations, the model does a reasonable job predicting DO and chlorophyll-a. This suggests that DO and chlorophyll-a are not a function of daily fluctuations in total phosphorus but more a function of long-term average concentrations, further suggesting that a "not to exceed at any time" total phosphorus standard may not be appropriate for this system. A monthly average or seasonal average may be more appropriate.

Evaluation of Reservoir Endpoints

Although there is some question as to whether wastewater treatment plants can achieve a 0.1 mg/l effluent total phosphorus concentration, the Panel has reservations that the urban and agricultural reductions of 80% can be attainable. In light of the general lack of enforceable legal mechanisms for controlling nonpoint source pollution, even if MS4s are considered to be nonpoint sources (whereas under the Clean Water Act they are point sources), it is doubtful whether an 80% nonpoint source reduction would be an achievable goal. Furthermore, even if we could "turn off" all inputs today, an 80% reduction is probably not achievable over the next ten to twenty years due to the build-up of phosphorus in the watershed. This level of nonpoint source reduction may not be practically achievable at all without major changes to the way individuals, farmers, towns, and corporations manage their land.

According to NJDEP's Stormwater Best Management Practice (BMP) Manual, the bioretention system (clearly NJDEP's most highly recommended BMP) can only achieve 60% reduction in total phosphorus, which is the highest phosphorus removal efficiency offered in the Manual for any recommended BMPs. Although stormwater management ordinances could be used to also decrease phosphorus loads, there are virtually no data demonstrating that the passage of any ordinance actual improves water quality. Even if the municipalities agree to pass phosphorus reduction ordinances, it will be difficult to enforce them on a wide-scale watershed basis. Reductions of 20% to 40% are recommended and have a much better chance of being achieved by the MS4s and the farmers.

Since the reservoir model has difficulties in simulating daily total phosphorus concentrations, the model should only be used to examine monthly averages. The success of the reservoir model is a function of the inputs the model receives from the river model. Since the river model has not demonstrated that it can accurately predict a daily time series for total phosphorus, any limitations developed by the reservoir model using inputs from the river model are suspect. Both models have shown that they tend to over-predict minimum concentrations. The use of this modeling system (i.e, the reservoir model in conjunction with the river model) to develop WLAs and LAs will likely result in overly stringent limitations due to the models' tendency to over-predict minimum concentrations. For example, if the model is predicting 0.14 mg/l concentrations when the observed is 0.07 mg/l, the point source loads and nonpoint source loads will be reduced beyond the needed level to bring the predicted concentration of 0.14 mg/l into compliance with state criteria.

Reference:

Najarian Associates, September 2000. "Water Quality Assessment of the Upper Passaic River Watershed and the Wanaque Reservoir." Prepared for New Jersey Department of Environmental Protection.